



①9 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑫ Offenlegungsschrift
⑩ DE 42 09 544 A 1

⑤1 Int. Cl.⁵:
H 03 M 7/30

②1 Aktenzeichen: P 42 09 544.1
②2 Anmeldetag: 24. 3. 92
④3 Offenlegungstag: 30. 9. 93

DE 42 09 544 A 1

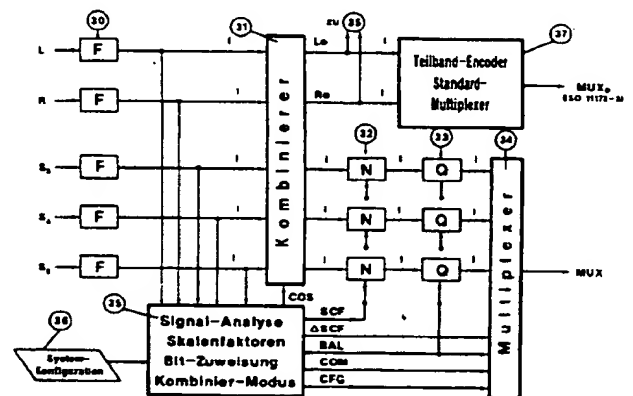
⑦1 Anmelder:
Institut für Rundfunktechnik GmbH, 80939 München,
DE
⑦4 Vertreter:
Konle, T., Dipl.-Ing., Pat.-Anw., 81247 München

⑦2 Erfinder:
Silzle, Andreas, 8000 München, DE; Stoll, Gerhard,
8051 Zolling, DE; Theile, Günther, 8191 Thanning, DE;
Link, Martin, 8000 München, DE

Prüfungsantrag gem. § 44 PatG ist gestellt

⑤4 Verfahren zum Übertragen oder Speichern digitalisierter, mehrkanaliger Tonsignale

⑤7 Zum Übertragen oder Speichern digitalisierter, mehrkanaliger Tonsignale die durch eine Vielzahl spektraler Teilbandsignale digital dargestellt sind werden encoderseitig die Teilbandsignale verschiedener Tonkanäle, jedoch gleicher Frequenzlage interkanalweise nach Maßgabe eines dynamischen Steuersignals kombiniert. Dieses Steuersignal wird durch eine an einem binauralen, psychoakustischen Modell orientierten Tonsignalanalyse von mehreren Tonkanälen gewonnen. Decoderseitig werden die Teilbandsignale verschiedener Tonkanäle, jedoch gleicher Frequenzlage interkanalweise in Abhängigkeit von einer Steuergröße dekombiniert, welche von dem dynamischen Steuersignal abgeleitet und mitübertragen bzw. mitgespeichert wird (Fig. 3).



DE 42 09 544 A 1

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

BUNDESDRUCKEREI 08. 93 308 039/141

21/50

Die Erfindung bezieht sich auf ein Verfahren gemäß dem Oberbegriff der nebengeordneten Ansprüche 1, 12 und 14.

Für die Übertragung und Speicherung digitaler Tonsignale ist es bekannt (DE 34 40 613 und DE 36 39 753), mit Hilfe einer Codierung in Teilbändern und der Anwendung eines psychoakustischen Modells eine Datenreduktion vorzunehmen. Dabei wird das digitalisierte, breitbandige Tonsignal durch eine Vielzahl spektraler Teilbandsignale dargestellt, wobei die Quantisierung der Abtastwerte in den einzelnen Teilbändern nach Maßgabe der Mithörschwelle des menschlichen Gehörs erfolgt. Der mit dieser "monofonen Codierung" erreichbare Datenreduktionsfaktor für sehr hochwertige Audiosignale hat bei 48 kHz Abtastfrequenz und 16 bit-linearer Quantisierung etwa den Wert 6, was einer Datenrate von $2 \times 128 \text{ kbit/s}$ pro stereofones Tonsignal entspricht.

Bekannt sind ferner Maßnahmen zur Steigerung des Datenreduktionsfaktors bei der Codierung zweikanaliger, stereofoner Signale, welche darauf beruhen, redundante Signalanteile im linken und rechten Kanal sowie solche Anteile zu unterdrücken, die bezüglich der stereofonen Wirkung irrelevant sind. Zwei unterschiedliche Ausführungsformen dieser "stereofonen Codierung" sind in der europäischen Patentanmeldung Nr. 90 00 338 sowie in "Proceedings of the ICASP 1990" "Perceptual Transform Coding of Wideband Stereo Signals" beschrieben, nämlich

- zeitweise Übertragung von Teilband-Summensignalen im oberen Frequenzbereich und blockweise Rekonstruktion der Teilband-Signalpegel im linken und rechten Kanal mit Hilfe von Skalenfaktoren, welche die maximalen Signalpegel in den Teilbändern der linken und rechten Tonkanäle repräsentieren, und

- Bildung von Summen-(M = MONO)- und Differenz-(S = SEITEN)-Signalen aus dem linken (L) und rechten (R) Tonsignal gemäß der Matrizierung $M = L + R$ und $S = L - R$, wobei nach Maßgabe der für das Summensignal M ermittelten Mithörschwelle die Decodierung der Summen- und Differenzsignale erfolgt.

Wird die getrennte Codierung in den Teilbändern des linken und rechten Tonkanals durch die stereofone Codierung in der Weise ergänzt, daß eine dieser Methoden oder eine Kombination aus beiden Methoden angewandt wird, so läßt sich der Datenreduktionsfaktor von dem Wert 6 auf etwa den Wert 8 steigern. Ist beispielsweise für die Übertragung von zwei unabhängigen Monosignalen mit derselben Qualität wie eine 16 bit-lineare Quantisierung eine Bitrate von 256 kbit/s erforderlich, so sind bei stereofoner Codierung für die Übertragung eines zweikanaligen Stereosignals derselben subjektiven Qualität nur 192 kbit/s erforderlich.

Demgegenüber besteht die Aufgabe der Erfindung darin, eine noch stärkere Datenreduktion bei zweikanaligen oder mehrkanaligen stereofonen Signalen zu erreichen.

Diese Aufgabe wird erfindungsgemäß durch die kennzeichnenden Merkmale der nebengeordneten Ansprüche 1, 12 oder 14 gelöst.

Vorteilhafte Ausgestaltungen und Weiterbildungen des erfindungsgemäßen Verfahrens ergeben sich aus

den Unteransprüchen.

Die Erfindung geht von der Überlegung aus, daß sich die Bitrate für ein teilbandcodiertes, stereofones Signal noch beträchtlich weiter senken läßt, wenn diverse Phänomene des räumlichen Hörens und bekannte Modelle der binauralen Signalverarbeitung im menschlichen Gehör berücksichtigt werden. Es zeigt sich nämlich, daß ein hoher Informationsanteil des stereofonen Signals irrelevant ist. Insbesondere kann das menschliche Gehör für bestimmte spektrale Komponenten und innerhalb bestimmter Zeitintervalle keine räumlichen Unterschiede wahrnehmen, d. h., innerhalb bestimmter Zeitintervalle ist für bestimmte Spektralbereiche eine vollständige Kanaltrennung nicht notwendig. Ferner kann infolge des Effektes der "Nachbarkanal-Verdeckung" die Quantisierung oder Unterdrückung einzelner Teilbandsignale nach Maßgabe der höchsten Nachbarkanal-Mithörschwelle erfolgen, wenn die zugehörige Gleichkanal-Mithörschwelle niedriger liegt. Allerdings muß für die Nutzung der Nachbarkanal-Verdeckung der Effekt einer reduzierten Verdeckung bei räumlich getrennten Maskierern ("Masking Level Difference", MLD) berücksichtigt und daher die Wiedergabeordnung definiert werden. Eine derartige Optimierung der stereofonen Codierung hat insbesondere bei zukünftigen digitalen Tonsystemen mit einer höheren Anzahl von Übertragungs- und Wiedergabekanälen eine besondere Bedeutung. Mit der Erhöhung der Anzahl der Tonkanäle lassen sich die Richtungsstabilität der Tondarstellung im Abbildungsbereich zwischen den vorderen Lautsprechern sowie die Möglichkeiten der räumlichen Tondarstellung steigern. Hierzu werden beispielsweise der linke und rechte Lautsprecher durch einen Center- sowie zwei Surround-Lautsprecher ergänzt, so daß drei weitere Tonübertragungskanäle erforderlich sind. Eine mit der Steigerung der Kanalzahl proportional wachsende Bitrate stellt in vielen Fällen eine zu hohe Belastung der Übertragungskapazität dar. Beispielsweise wäre bereits eine Verdopplung der Bitrate für zusätzliche Center- und Surroundkanäle im zukünftigen digitalen Hörfunk nicht akzeptabel, wenn die Anzahl der Programme entsprechend halbiert werden müßte.

Es wird daher speziell angestrebt, die Bitrate für fünfkanalige, stereofone Tonsignale (3 Frontsignale L (Links), C (Center), R (Rechts) und 2 Surroundsignale LS (Links-Surround), RS (Rechts-Surround); abgekürzt mit "312-Stereosignale" von $5 \times (192:2) = 480 \text{ kbit/s}$ auf etwa 300 kbit/s zu verringern. Eine wichtige Bedingung für die Einführung mehrkanaliger Tonübertragungssysteme ist in vielen Fällen die Kompatibilität mit vorhandenen digitalen Zweikanal-Stereoempfängern. Da die Übertragungskapazität oft zu begrenzt ist, um entsprechend dem Simulcast-Prinzip neben dem konventionellen Zweikanal-Signal gemäß ISO/EMPEC-Standard 11172-3 ein vollständiges Mehrkanal-Signal zu übertragen, muß mit Hilfe von Kompatibilitätsmatrizierungen vor der Übertragung aus dem Mehrkanal-Stereosignal ein Zweikanal-Stereo-Basissignal L_o , R_o gewonnen werden, welches für die Wiedergabe mit Zweikanal-Stereoempfängern geeignet ist. Für eine Kompatibilitätsmatrizierung für fünfkanalige (3/2)-Stereosignale kann beispielsweise folgendes Gleichungssystem vorgesehen werden:

$$T1 = L + xC + yLS = L_o \quad (1)$$

$$T2 = R + xC + yRS = R_o \quad (2)$$

$$T3 = xC \quad (3)$$

$$T4 = yLS \quad (4)$$

$$T5 = yRS \quad (5)$$

wobei x und y Koeffizienten im Bereich von beispielsweise 0,5 bis 0,8 sind.

Die Übertragungssignale T1 und T2 bilden das Zweikanal-Stereo-Basissignalspaket Lo/Ro, während die zusätzlichen Übertragungssignale T3, T4, T5 die für die Dematrizierungen notwendige Information enthalten. Sollen die ursprünglichen fünfkanaligen Stereosignale L, C, R, LS, RS decoderseitig wieder vollständig rekonstruiert werden, so lautet die Dematrizierungsvorschrift:

$$L' = T1 - T3 - T4 \quad (6)$$

$$R' = T2 - T3 - T5 \quad (7)$$

$$C' = T3/x \quad (8)$$

$$LS' = T4/y \quad (9)$$

$$RS' = T5/y \quad (10)$$

Eine vollständige Rekonstruktion der ursprünglichen fünfkanaligen Stereosignale ist jedoch nicht notwendig, da — wie schon erwähnt — das Gehör innerhalb bestimmter Zeitintervalle für bestimmte Spektralbereiche eine unvollständige Kanaltrennung toleriert. Dementsprechend können bestimmte Teilbandsignale der Übertragungssignale T3, T4 und T5 in den Gleichungen (3), (4) und (5) zeitweise und signalabhängig zu Null gesetzt werden (sogenannte "Punktierung"). Eine weitere Überlegung besteht daher darin, die Datenrate für das Zusatzsignalspaket T3/T4/T5 so weit wie möglich zu senken, indem ausschließlich die für die räumliche Wahrnehmung relevanten Signalanteile extrahiert werden. Die Codierung der Stereo-Basissignale Lo und Ro bleibt hiervon unberührt, so daß Kompatibilität mit bestehenden Zweikanal-Decodern gewährleistet ist. Anstelle einer mehrkanaligen Tondarstellung lassen sich auch mehrsprachige, einfache Programme oder Kommentarkanäle vorsehen, insbesondere für die Anwendung in zukünftigen Fernsehsystemen mit digitalem Ton. In diesem Fall wird keine Matrizierung vorgesehen.

Die Erfindung wird nachstehend anhand von Ausführungsbeispielen in den Zeichnungen näher erläutert. Es zeigt:

Fig. 1 ein Grundschema eines Encoders zur encoderseitigen Durchführung des erfindungsgemäßen Verfahrens;

Fig. 2 ein Grundschema eines Decoders zur decoderseitigen Durchführung des erfindungsgemäßen Verfahrens;

Fig. 3 und 5 jeweils ein Grundschema eines kompatiblen Encoders für fünfkanalige, stereofone Tonsignale;

Fig. 4 und 6 jeweils ein Grundschema eines kompatiblen Decoders für fünfkanalige, stereofone Tonsignale und

Fig. 7 ein Modell der binauralen Signalverarbeitung des menschlichen Gehörs.

Bei der Erfindung werden neben der Redundanz vor allem die Grenzen des Gehörs bei der räumlichen Analyse des Schallfeldes im Sinne einer stereofonen Codie-

rung möglichst vollständig genutzt, indem die Bitzuweisung für individuelle Teilbänder einzelner Tonkanäle blockweise nach Maßgabe einer gemeinsamen Analyse der stereofonen Signale aller Tonkanäle erfolgt. Hierbei wird die Quantisierung nicht nur unter Berücksichtigung der Nachbarkanal-Verdeckung festgelegt, sondern darüber hinaus unter Berücksichtigung der subjektiv hinreichenden Kanaltrennung, wie im folgenden näher dargelegt werden soll.

Eine signalabhängig gesteuerte, gehörangepaßte, blockweise Reduktion der Kanaltrennung individueller Teilbandsignale führt zu entsprechender Datenreduktion, wenn die im Encoder zu "mono" gesetzten Teilbandsignale nur in einem der Kanäle übertragen werden und im Decoder ein geeigneter Austausch einzelner Teilbandsignale möglich ist. Die betreffenden Teilbänder der restlichen Tonkanäle werden nicht übertragen. Hierzu sind im Encoder und im Decoder Netzwerke ("Kombinierer" bzw. "Dekombinierer") vorgesehen, die nach Maßgabe von mitübertragenen Steuerinformationen, die aus einer encoderseitigen Signalanalyse gebildet werden und individuelle Teilbandsignale in demselben Frequenzbereich einzelnen Tonkanälen zuweisen oder interkanalweise verknüpfen. Wie im folgenden noch erläutert wird, ist es besonders vorteilhaft, die teilband- und blockweise Reduktion der Kanaltrennung auf die Abtastwerte zu beschränken und die encoderseitig gewonnenen, als Nebeninformation übertragenen, ursprünglichen Skalenfaktoren der einzelnen Tonkanäle für die Rückgewinnung der Teilbandsignale einzusetzen.

Ein beispielhaftes Grundschema eines Encoders und Decoders für n Tonkanäle geht aus den Fig. 1 und 2 hervor. Im Encoder gemäß Fig. 1 werden die digitalisierten, linear-quantisierten Tonsignale S1 bis Sn einer Teilbandfilterbank 10 zugeführt und jeweils in i Teilbänder zerlegt. Die resultierenden nxi Teilbandsignale durchlaufen nacheinander einen Kombinierer 11, einen Skalierer 12 sowie einen Quantisierer 13 und werden zusammen mit den in einer Analysestufe 15 gewonnenen Nebeninformationen BAL (Bit-Zuweisungs- bzw. -Allokationsinformation), COM (Kombinier-Modus-Information) und SCF (Skalenfaktoren) in einem Multiplexer 14 zu einem Multiplexsignal MUX zusammengefaßt. Dem Kombinierer 11 wird von der Analysestufe 15 ein Steuersignal COS zugeführt, welches im Kombinierer 11 eine interkanalweise Zusammenfassung individueller Teilbandsignale steuert. In dem Skalierer 12 werden die Abtastwerte nach Maßgabe der Skalenfaktorinformation SCF normiert, während in dem Quantisierer 13 nach Maßgabe der Bit-Zuweisungsinformation BAL die Quantisierung der normierten Abtastwerte erfolgt. Wie die Berechnung aller Nebeninformationen zur Steuerung der Stufen 11, 12, 13 durch die Analysestufe 15 erfolgt, wird später noch im einzelnen dargelegt.

In dem Decoder gemäß Fig. 2 werden nach der Zerlegung des empfangenen Datenstroms MUX in einem dem Demultiplexer 24 nachgeschalteten Dequantisierer 23 die Abtastwerte der Teilbandsignale mit Hilfe der übertragenen Bit-Zuweisungsinformation BAL in normierter Form zurückgewonnen. Die zurückgewonnenen Abtastwerte werden nach Maßgabe der Kombinier-Modus-Information COM in einem Dekombinierer 21 auf n Tonkanäle teilbandweise aufgeteilt. Erst danach erfolgt eine Multiplikation der Teilbandsignale mit den Skalenfaktoren SCF in Multiplizierstufen 22. Nach der inversen Filterung der entnormierten Teilbandsignale in

den Synthesefilterbänken 20 liegen an den Ausgängen des Decoders gemäß Fig. 2 die breitbandigen, digitalisierten Tonsignale S1' bis Sn' vor.

Eine besondere Ausführungsform des erfindungsge-
mäßigen Verfahrens ist in den Fig. 3 und 4 dargestellt.
Berücksichtigt ist hier die eingangs beschriebene Forde-
rung der Kompatibilität mit bestehenden Teilband-Codierern gemäß dem digitalen Standard ISO 11 172-3 so-
wie die Forderung nach flexibler Nutzung der Tonüber-
tragungskanäle. Ausgegangen wird im Falle des Enco-
ders nach Fig. 3 von fünfkanaligen, stereofonen, digitali-
sierten Tonsignalen, nämlich 3 Frontsignalen Links (L),
Rechts (R) und Center (C) sowie zwei Surroundsignalen
Links-Surround (LS) und Rechts-Surround (RS). Aus
Kompatibilitätsgründen zu bestehenden Standard-Teil-
band-Codierern werden die in Teilbandfiltern 30 in je-
weils i Teilbänder gefilterten Frontsignale L und R in
einem Kombiniierer 31 beispielsweise gemäß vorstehen-
der Gleichungen (1) und (2) zu einem Zweikanal-Stereo-
Basissignalpaket Lo'/Ro' matriziert, welches in einer als
"Teilband-Encoder und Standard-Multiplexer bezeich-
neten Stufe 37 in bekannter Weise teilband-codiert und
in einen standardisierten Bitstrom gemäß ISO-Standard
11 172-3 unter Bildung eines kompatiblen Stereosignals
MUXo umgewandelt wird.

Die in Teilbandstufen 30 in jeweils i Teilbänder gefil-
terten Center- und Surround-Signale C, LS, RS werden
in dem Kombiniierer 31 beispielsweise gemäß Gleichun-
gen (3) bis (5) zu einem teilbandgefilterten Zusatzsignal-
paket T3/T4/T5 matriziert. Dieses Zusatzsignalpaket
T3/T4/T5 wird in gleicher Weise wie in Fig. 1 einer Skali-
erung (Skalierer 32), einer Quantisierung (Quantisierer
33) und einer Multiplexbildung (Multiplexer 34) unter-
zogen, wobei die Steuerung der Stufen 31, 32, 33 und 34
wiederum durch eine Analysestufe 35 erfolgt. Im Ver-
gleich zur Analysestufe 15 gemäß Fig. 1 erzeugt die
Analysestufe 35 gemäß Fig. 3 zusätzlich eine Differenz-
Skalenfaktorinformation Δ SCF sowie eine Konfigura-
tionsinformation CFG, welche in das Multiplexsignal
MUX des Multiplexers 34 eingefügt werden. Die Analy-
sestufe 35 erhält von einer anwenderbedienten System-
Konfigurations-Eingabe 36 eine applikationsabhängige
Information über die Art der gewünschten Verarbeitung
der Eingangssignale, beispielsweise eine Fünfkana-
l-Datenreduktion unter Berücksichtigung der stereo-
fönen Irrelevanz aller Kanäle (Surround-Tondarstel-
lung) oder eine Codierung der Eingangssignale C, LS,
RS unabhängig von den Frontsignalen L, R, beispiels-
weise für einen mehrsprachigen Fernsehton. LS, RS, C
sind im letztgenannten Beispielsfall keine Surround-Ka-
näle, sondern alternative Sprachkanäle, wie an Hand
von Fig. 5 noch erläutert wird. Zusätzlich zu den n teil-
bandgefilterten Eingangssignalen erhält die Analysestu-
fe 35 die teilbandgefilterten Stereo-Basissignale Lo, Ro.

Die vorgenannten Eingangssignale werden in der
Analysestufe 35 entsprechend dem dort verwendeten
psychoakustischen Modell und den daraus abgeleiteten,
nachfolgend beschriebenen Kriterien analysiert. Ergibt
die Analyse, daß eine Kanaltrennung subjektiv nicht
erforderlich ist, so wird eine Aufhebung der Kanaltren-
nung dadurch erreicht, daß mit Hilfe der Bit-Zuwei-
sungsinformation BAL bestimmte Teilbandsignale des
Zusatzsignalpaketes C/LS/RS im Quantisierer 33 block-
weise zu Null gesetzt werden. Die Bildung des Zusatz-
Multiplexsignals MUX im Multiplexer 34 umfaßt auch
quellenbezogene Fehlerschutzmaßnahmen. Die Diffe-
renz-Skalenfaktorinformation Δ SCF besteht aus der
Differenz zwischen den Skalenfaktoren SCFo der Teil-

bänder der Stereo-Basissignale Lo, Ro und den Skalenfak-
toren SCF der Teilbänder der Zusatzsignale C, LS, RS
und ermöglichen daher eine bitsparendere Codierung
der Skalenfaktoren SCF. Anstelle einer expliziten Über-
tragung der Skalenfaktoren SCF im Multiplexsignal
MUX gemäß Fig. 1 werden bei dem Encoder gemäß
Fig. 3 nur die Differenzwerte Δ SCF übertragen.

Die resultierenden Bitströme MUX und MUXo des
Encoders gemäß Fig. 3 werden in beliebiger Weise in
getrennten Datenpaketen übertragen oder gespeichert.
Am Ort des Decoders gemäß Fig. 4 muß gewährleistet
sein, daß beide Bitströme MUX und MUXo zeitgleich
verfügbar sind. Es ist ferner möglich, die resultierenden
Bitströme MUX und MUXo in einem übergeordneten
Multiplex zu übertragen bzw. zu speichern. Hierzu wird
beispielsweise, wie in Fig. 5 gezeigt ist, der Bitstrom
MUX der Stufe 37 zugeführt, welche den Bitstrom
MUX in geeigneter Weise in den Bitstrom MUXo unter
Bildung des kombinierten Bitstroms MUXo* einfügt.
Dies kann beispielsweise bei einem Datenstrom gemäß
ISO-Standard 11 172-3 dadurch erfolgen, daß in jedem
Audio-Datenrahmen innerhalb des Feldes, das für pro-
grammbezogene Daten reserviert ist, die MUX-Daten
eingefügt werden.

In dem Decoder gemäß Fig. 4 werden aus dem Bits-
trom MUXo in einer als "Standard-Demultiplexer und
Teilband-Decoder" bezeichneten Stufe 47 in bekannter
Weise die Teilbandsignale des Zweikanal-Stereo-Basis-
signalpaketes Lo'/Ro' zurückgewonnen. Ferner werden
aus dem Bitstrom MUX nach erfolgter Demultiplexbil-
dung in einem Demultiplexer 44 die Teilbandsignale des
Zusatzsignalpaketes mit Hilfe von Dequantisierern 43
und Multiplizierern 42 zurückgewonnen. Die hierzu
notwendigen Skalenfaktoren SCF der Teilbandsignale
C, LS, RS werden in einer Kontroll-Einheit 45 aus den
von dem Demultiplexer 44 gelieferten Differenz-Ska-
lenfaktorinformationen Δ SCF und den von der Stufe 47
gelieferten Skalenfaktoren SCFo der Teilband-Basissig-
nale Lo, Ro errechnet. Wie im Falle von Fig. 2 werden
die Dequantisierer 43 von der Bit-Zuweisungsinforma-
tion BAL und die Multiplizierer 42 von den in der Kon-
troll-Einheit 45 errechneten Skalenfaktoren SCF ge-
steuert.

Die decodierten Teilbandsignale Lo', Ro' am Aus-
gang der Stufe 47 sowie die decodierten Teilbandsignale
T3', T4' und T5' am Ausgang der Multiplizierer 42 ge-
langen an den Dekombiniierer 41, wo die bereits er-
wähnte Kanalzuteilung oder die interkanalweise Ver-
knüpfung der Teilbandsignale erfolgt, und zwar unter
der Steuerung der Kontroll-Einheit 45, welche das Steu-
ersignal COS aus den vom Demultiplexer 44 gelieferten
Zusatzinformationen COM und CFG sowie aus einem
Eingabesignal EIN errechnet. Das Eingabesignal EIN
wird von einem Anwender über eine Wiedergabe-Kon-
figurations-Eingabe 46 eingegeben; es repräsentiert die
gewünschte Abhörsituation sowie — im Falle des mehr-
sprachigen Fernsehtons — die gewünschte Wiederga-
besprache. Desweiteren kann der Anwender über die
Eingabe 46 Vorgaben über Pegelbalance und Abhördy-
namik in die Kontroll-Einheit 45 eingeben, welche nach
Maßgabe dieser Eingaben die dem Dekombiniierer 41
nachgeschalteten Gewichtungsstufen 48 steuert, welche
die Teilbandsignale an den Ausgängen des Dekombinie-
rers 41 blockweise mit entsprechenden Gewichtungs-
faktoren GEW multipliziert. Nach der anschließenden
inversen Filterung in den Synthesefilterbänken 40 liegen
an den Ausgängen des Decoders gemäß Fig. 4 die breit-
bandigen digitalisierten Tonsignale L', R', S3', S4' und

55' vor. In dem Decoder gemäß Fig. 6 werden im Unterschied zum Decoder gemäß Fig. 4 aus dem kombinierten Datenstrom MUXo* in der Stufe 47 der Datenstrom MUX sowie die Skalenfaktoren SCFo der Teilband-Basis-signale Lo, Ro abgetrennt und dem Demultiplexer 44 zugeführt. Die direkte Zuführung von SCFo von der Stufe 47 zu der Kontroll-Einheit 45 (Fig. 4) entfällt; vielmehr werden in Fig. 6 die Skalenfaktoren SCFo von dem Demultiplexer 44 der Kontroll-Einheit 45 zugeführt. Im übrigen ist der Decoder gemäß Fig. 6 identisch mit dem Decoder gemäß Fig. 4.

Für die nähere Erläuterung des erfindungsgemäßen Verfahrens an den Ausführungsbeispielen gemäß Fig. 3 bis 6 werden nachfolgend zunächst die der stereofonen Codierung zugrunde liegenden Höreigenschaften dargestellt und darauf basierend die eingesetzten Maßnahmen für die Datenreduktion. Weitere Ausgestaltungen, insbesondere in Hinblick auf eine flexible Nutzung der Übertragungskanäle, werden abschließend beschrieben.

Modell der binauralen Signalverarbeitung

Für das räumliche Hören haben Unterschiede zwischen den Schallsignalen an beiden Ohren (interaurale Signaldifferenzen) eine hervorragende Bedeutung. In der Literatur werden mehrere Modelle unterschieden, die bereits eine weitreichende Beschreibung binauraler Effekte zulassen. Allerdings nehmen die Kenntnisse über die zugrundeliegenden physiologischen Verarbeitungsmechanismen immer mehr ab, je weiter man in die zentralen neuronalen Bereiche vordringt. Aus diesem Grunde sind die peripheren Stufen binauraler Modelle zumindest in ihrer grundsätzlichen Konzeption einheitlich, während die zentralen Modellbereiche entweder auf eng umgrenzte psychoakustische Effekte spezialisiert sind oder nur unscharfe Aussagen ermöglichen. Ein Modell der binauralen Signalverarbeitung zeigt Fig. 7. Die relativ genaue Beschreibung der peripheren Stufe durch die Bausteine A, B, C, D sowie recht allgemeine Annahmen über die Funktionsweise der zentralen Verarbeitungsstufe ("Binauraler Processor") erlauben eine grobe Erfassung der "binauralen Information". Das Modell gemäß Fig. 7 enthält folgende Stufen:

Außenohr-Filter:

Die Wirkung des Außenohres (Ohrmuschel, Kopf, Rumpf) auf das linke und rechte Ohrsignal ist durch Außenohr-Übertragungsfunktionen (Außenohr-Filter) beschrieben. Hierdurch erfolgt eine räumliche Zerlegung sowie eine spektrale und zeitliche Filterung des Quellensignals im Sinne einer "räumlichen Codierung". Die Wirkung ist unterhalb etwa 150 Hz vernachlässigbar, d. h. entsprechend tieffrequente Quellensignale sind hinsichtlich der binauralen Auswertung irrelevant.

Periphere Stufe A:

Es gilt nicht nur für die monaurale, sondern ebenso für die binaurale Signalverarbeitung als erwiesen, daß die Auswertung jedes Ohrsignals frequenzselektiv geschieht. Das Gehör zerlegt die Ohrsignale in schmalbandige Anteile f_1, f_2, \dots, f_n . Die Analysebandbreite der betreffenden Schmalband-Filter entspricht den Frequenzgruppenbreiten, die bei monauralen Experimenten gemessen werden, d. h. die auch physiologisch nachgewiesene periphere Frequenzselektivität wird sowohl bei monauraler als auch bei binauraler Signalverarbeitung wirksam.

Periphere Stufe B:

Ebenfalls physiologisch und psychoakustisch begründet

ist der folgende periphere Modellbaustein B zur Nachbildung der neuronalen Eigenschaften. Die Ausgangssignale der Haarzellen, welche in der Cochlea die mechanische Schwingungsform entlang der Basilarmembran abtasten, sind Folgen von Nervenimpulsen (Spikes), deren zeitliche Dichte in dem betreffenden Frequenzband mit wachsendem Signalpegel zunimmt ("Pulsfrequenzmodulation"). Zur Beschreibung der für das Richtungs-hören wichtigen Eigenschaften der Nervenimpulsfolgen ist es in erster Näherung ausreichend, ein einfaches Receptor-Neuronen-Modell aus Halbweggleichrichtung und anschließender Tiefpaßfilterung für jede Frequenzgruppe vorzusehen. Die Halbweggleichrichtung berücksichtigt, daß die Haarzellen nur auf eine Halbwelle der Basilarmembranschwingung ansprechen. Der Tiefpaß 1. Ordnung mit einer Grenzfrequenz von ca. 800 Hz dient zur Gewinnung der Hüllkurve und berücksichtigt, daß die Aktivitätswahrscheinlichkeit der Neuronen mit steigender Signalfrequenz mehr und mehr der Hüllkurve und weniger der Feinstruktur der Basilarmembranbewegung folgt, im Bereich tiefer Frequenzen dagegen Proportionalität vorliegt.

Periphere Stufe C:

Die aus einem Dividierglied und einem rückgekoppelten RC-Glied dargestellte, nichtlineare Stufe C bewirkt für stationäre Signale eine Dynamikkompression, während schnelle Änderungen des Eingangssignals weitgehend linear übertragen werden. Hierdurch werden zeitliche Verdeckungseffekte (insbesondere Nachverdeckung) nachgebildet; ferner ist die Stufe C ebenso für wichtige binaurale Zeiteffekte primär verantwortlich (z. B. "Gesetz der ersten Wellenfront", "Franssen-Effekt"). Die Zeitkonstante für den Abfall der Hüllkurve liegt im Durchschnitt etwa bei 50 ms, dagegen für den Anstieg bei 1,25 ms = 1/800 Hz. Hierdurch ergibt sich eine hohe Informationsreduktion.

Periphere Stufe D:

Die Ausgangssignale der Stufe C werden in der peripheren Stufe D einer Zeit-Frequenztransformation (z. B. Fast-Fourier-Transformation) unterzogen, die die spektralen Eigenschaften der Hüllkurve darstellt. Sie ändern sich relativ langsam, was den Änderungen der Aktivitätszustände im akustischen Cortex entspricht. Die frequenzselektiven Ausgangssignale liegen als Muster der binauralen Verarbeitungsstufe (binauraler Processor) zeitgleich vor.

Binauraler Processor:

Für die frequenzselektive Verarbeitung im binauralen Processor wird angenommen:

1. Die Auswertung erfolgt durch Mustererkennungsmechanismen, d. h. durch Vergleich des aktuellen Signalmusters mit einem Vorrat gespeicherter (erlernter) Muster. Dies ist ein Prozeß, der ein aktuelles Signalmuster mit einem gespeicherten Muster verknüpft, selbst dann, wenn nur Teile des gespeicherten Musters im aktuellen Signalmuster enthalten sind oder wenn nur ein unvollständiges aktuelles Signalmuster vorliegt.
2. Voraussetzung für die Erkennung räumlicher Eigenschaften ist Kohärenz individueller kontralateraler Hüllkurven-Signale, ("binaural envelope coherence"), d. h. individuelle Hüllkurvensignale vom linken und rechten Ohr mit gleicher Frequenzlage stellen Musterelemente dar, die hinsichtlich ihrer binauralen Information ausgewertet werden.
3. Die so gewonnenen binauralen Informationen werden den entsprechenden, in einem separaten

Mustererkennungsprozeß gewonnen, monauralen Informationen zugeordnet (orts- und gestaltsbestimmende Mustererkennung).

Selektion der stereofonen Information

In Hinblick auf die stereofone Codierung macht das Modell gemäß Fig. 7 zunächst die wichtige Aussage, daß die Selektion der stereofonen Information möglich ist, wenn die Auswertung frequenzselektiv erfolgt, d. h. wenn eine Zerlegung in (möglichst frequenzgruppenbreite) Teilbänder vorgesehen wird. Die Eigenschaften der frequenzselektiven Signale an den Ausgängen der Stufen D sowie Eigenschaften der Mustererkennung geben Hinweise für die Codierung der stereofonen Satzsignale in den Teilbändern.

Eine hervorragende Bedeutung haben demnach die Hüllkurven aller Teilbandsignale. Im Bereich höherer Frequenzen (oberhalb etwa 1,5 kHz) lassen sie keine Auswertung der interauralen Phasendifferenzen zu. Dies ist sinnvoll, da oberhalb etwa 1,5 kHz die interaurale Phasendifferenz kein eindeutiges Richtungsmerkmal mehr darstellt. Lateralisationsexperimente zeigen, daß im gesamten Hörfrequenzbereich (ab etwa 150 Hz) interaurale Pegel- und Zeitdifferenzen der Hüllkurven zu einer seitlichen Auslenkung des Hörereignisses führen, dies aber bei interauralen Zeitdifferenzen der Signaleinstrukturen nur im Frequenzbereich bis ca. 1,5 kHz geschieht. Mit amplitudenmoduliertem Hochpaßrauschen findet man selbst dann ein fusioniertes Hörereignis, wenn interaural unkorreliertes Rauschen verwendet wird, sofern nur beide Ohrsignale die gleichen Hüllkurven aufweisen.

Diese Teilband-Hüllkurvenauswertung bedeutet bezüglich der stereofonen Codierung, daß die stereofone Information der oberen Teilbandsignale allein durch ihre Hüllkurven repräsentiert ist, und daß demzufolge die stereofonen Signale bei höheren Frequenzen zusammengefaßt werden dürfen, wenn vorher die Hüllkurven der oberen Teilbandsignale extrahiert worden sind. Diese stereofone Information ist geeignet, im Decoder eine subjektiv ausreichende Kanaltrennung wiederherzustellen, indem die monofonen Teilbandsignale mit den ursprünglichen Hüllkurven moduliert werden. Hierdurch entstehen stereofone Signale, die in höheren Frequenzlagen sich nur bezüglich ihrer Hüllkurven, nicht aber bezüglich der Signaleinstruktur unterscheiden, und die dennoch ein unbeeinträchtigtes stereofones Klangbild sicherstellen. Voraussetzung dafür ist die hinreichend genaue Reproduktion der Teilband-Hüllkurven. Es hat sich bereits erwiesen, daß sogar bei grober zeitlicher Auflösung die resultierenden Qualitätsbeeinträchtigungen relativ gering sind. Sie lassen sich vollkommen vermeiden, wenn die Reduktion der Kanaltrennung nur zeitweise und insbesondere im obersten Frequenzbereich häufiger erfolgt als darunter.

Im folgenden sollen die in dem binauralen Processor auftretenden zeitlichen Effekte und deren Auswertung bei der stereofonen Codierung betrachtet werden. Das zeitliche Verhalten der Signale an den Ausgängen der Stufe D sowie der Prozeß der Mustererkennung bewirken Einschränkungen der binauralen Zeitanalyse, die sich durch drei Effekte ausdrücken:

— Der "Franssen-Effekt" besagt, daß zwei räumlich getrennte Schallereignisse ein einziges Hörereignis am Ort desjenigen Lautsprechers erzeugen, welcher die Einschwingvorgänge des einen Schaller-

eignisses abstrahlt. Der eher stationäre Teil des anderen Schallereignisses ist für die Richtungszuweisung irrelevant.

— Das "Gesetz der ersten Wellenfront" besagt, daß zwei räumlich getrennte Schallereignisse ein einziges Hörereignis am Ort desjenigen Lautsprechers erzeugen, welcher das voreilende (impulsbehaftete) Tonsignal abstrahlt. Das verzögerte Tonsignal ist für die Richtungszuweisung irrelevant.

— Der "Trägheits-Effekt" besagt, daß Richtungsänderungen bei vorwiegend stationären Schallen erst ab einer bestimmten Änderungsdauer wahrscheinlich sind. Ein Stereosignal kann deshalb beispielsweise kurzzeitig ("weich") auf Mono geschaltet werden, ohne Störungen des stereofonen Klangbildes zu verursachen.

Es läßt sich als gemeinsames Merkmal der gesamten drei Effekte feststellen, daß eine Reihe kurzzeitiger Schallereignisse irrelevant ist bezüglich räumlicher Wahrnehmbarkeit und daher als räumliche Information entbehrlich ist. Diese irrelevanten Schallereignisse sind sogar länger, je schmalbandiger sie sind. Die zeitliche Lage dieser irrelevanten Schallereignisse bzw. Tonsignal-Abschnitte läßt sich anhand der Teilbandsignale-Hüllkurvenverläufe ermitteln; in erster Näherung liegen sie direkt hinter ihren Anstiegsflanken.

Merkmale des Verfahrens

Der Kombiniierer 31 im Encoder (Fig. 3 und 5) sowie der Dekombiniierer 41 im Decoder (Fig. 4 und 6) ermöglichen es, teilband- und blockweise beliebige Matrizierungen und Dematrizierungen vorzunehmen bzw. aufzuheben. Hieraus ergeben sich einerseits vorteilhafte Möglichkeiten der Datenreduktion und andererseits Möglichkeiten der alternativen Nutzung der Tonübertragungskanäle, beispielsweise Übertragung alternativer Sprachen, abhängig von der Wahl der System-Konfiguration mittels der Eingabe 36 (Fig. 3 und 5).

System-Konfiguration A

Diese System-Konfiguration repräsentiert die Datenreduktion für fünfkanalige stereofone Signale, d. h., es ist

$$\begin{aligned} S3 &= C \\ S4 &= LS \\ S5 &= RS. \end{aligned}$$

Der Kombiniierer 31 im Encoder gemäß Fig. 3 und 5 matriziert die Signale L, R, C', LS', RS' gemäß den Gleichungen (1) bis (5), während im Dekombiniierer 41 des Decoders gemäß Fig. 4 und 6 die Dematrizierung gemäß Gleichungen (6) bis (10) erfolgt. Die Datenreduktion geschieht durch folgende Maßnahmen:

1. Dynamisches Übersprechen

Eine hohe Datenreduktion für das Zusatzsignalpaket T3/T4/T5 ist erreichbar, indem nach Maßgabe einer geeigneten Signalanalyse in der Analysestufe 35 in individuellen Teilbändern die Dematrizierung blockweise aufgehoben wird. Dies geschieht mit Hilfe einer blockweisen sogenannten "Punktierung" von Teilbandsignalen des Zusatzsignalpaketes T3/T4/T5 (Gleichungen (3), (4), (5)). Bei der Punktierung werden die Abtastwerte des entsprechenden Teilbandsignals encoderseitig im Kombiniierer 31 und/oder Quantisierer 33 nach Maßgabe des Steuersignals COS und/oder der Bit-Zuwei-

sungsinformation BAL zu Null gesetzt. Decoderseitig werden die fehlenden Abtastwerte in den Teilbändern durch übertragene Abtastwerte in den Teilbändern anderer Tonkanäle, jedoch gleicher Frequenzlage nach Maßgabe des Steuersignals COM substituiert. Die substituierten Abtastwerte werden anschließend so in den Stufen 48 gewichtet, daß eine Anpassung an die ursprünglichen Teilbandsignalpegel-Verhältnisse erfolgt. Zur Durchführung dieser Anpassung wird der Gewichtungsfaktor GEW (Fig. 3 und 5) in der Kontroll-Einheit 45 aus den Skalenfaktoren SCF und dem Kombiniert-Modus COM errechnet, beispielsweise mittels Interpolation aufeinanderfolgender Skalenfaktoren, um "harte" Pegelsprünge zu vermeiden.

Die Regeln für die Anwendung der Punktierung ergeben sich aus verschiedenen Gehöreigenschaften, die im vorstehenden im Einzelnen erläutert worden sind. Grundsätzlich erfolgt die Punktierung in einer Weise, daß die für die binaurale Signalverarbeitung relevante Information im Zusatzsignalpaket T3/T4/T5 erhalten bleibt. Somit lassen sich die im Stereo-Basissignalpaket Lo/Ro zusammengefaßten Einzelsignale decoderseitig den ursprünglichen Tonkanälen so weit zuordnen, daß das stereofone Klangbild subjektiv wieder hergestellt wird.

Im Einzelnen werden die Tonsignale in der Analysestufe 35 hinsichtlich folgender Punktierungs-Kriterien analysiert:

— Teilbandsignal-Dynamik

Gemäß dem "Franssen-Effekt" und dem "Gesetz der 1. Wellenfront" können Teilbandsignalblöcke punktiert werden, welche einem Transienten folgen und sich stationär oder quasi-stationär verhalten.

— Hüllkurven-Auswertung

Die Auswertung der Hüllkurven der Teilbandsignale im binauralen Modell gemäß Fig. 7 erlaubt es, Punktierungen in Detailbändern höherer Frequenzlage häufiger vorzunehmen als in den Teilbändern tieferer Frequenzlage.

— Signalenergie-Verhältnisse

Die Energie eines Teilbandes läßt sich aus dessen Skalenfaktoren SCF ableiten. Die Aufsummierung der Signalenergien sämtlicher Teilbänder ergibt die Signalenergie des gesamten Tonsignals. Durch Vergleich der Signalenergie der einzelnen Teilbänder mit der Signalenergie des gesamten Tonsignals läßt sich ein weiteres Kriterium für die Punktierung ableiten. Und zwar können Punktierungen in denjenigen Teilbändern häufiger erfolgen, in denen die Signalenergie bewegten auf die Signalenergie des gesamten Tonsignals relativ gering ist. Unter "relativ gering" versteht man beispielsweise ein Verhältnis von 1/i, wobei i die Anzahl der Teilbänder je Tonkanal ist.

— Trägheits-Effekt des menschlichen Gehörs

Wie vorstehend bereits erwähnt, versteht man unter dem "Trägheits-Effekt" des menschlichen Gehörs, daß Richtungsänderungen bei vorwiegend stationären Schallen erst ab einer bestimmten Änderungsdauer wahrscheinlich sind. Daraus resultieren zwei Möglichkeiten für eine Punktierung:

a) Wird nur eine kurzzeitige Richtungsänderung im Decoder dedektiert, können die dafür verantwortlichen Teilbänder dennoch punktiert werden.

b) Wenn Punktierungen nur kurzzeitig erfolgen, beispielsweise aus Gründen einer kurzzei-

tigen Überlastung des Übertragungskanals für den Datenstrom MUx, kann das Gehör die dadurch verursachten Störungen in der räumlichen Abbildung wahrnehmen.

— Teilweise Nachbarkanal-Verdeckung

Der bereits erwähnte Effekt der "Nachbarkanal-Verdeckung" kann nicht nur — wie im Folgenden noch im Einzelnen ausgeführt wird — zur Festlegung der Quantisierung oder zur Unterdrückung einzelner Teilbandsignale ausgenutzt werden, sondern auch als weiteres Punktierungs-Kriterium herangezogen werden. Beispielsweise können solche Teilbänder, die nur teilweise bzw. nicht vollständig durch Nachbarkanäle verdeckt werden, eher punktiert werden als solche Teilbänder, die keiner Nachbarkanal-Verdeckung unterliegen.

— Redundanz-Ausnutzung

Aufgrund der Kompatibilitätsmatrizierung gemäß Gleichungen (1) bis (5) sind zeitweise einzelne Teilbandsignale von T1 und/oder T2 mit entsprechenden Teilbandsignalen gleicher Frequenzlage von T3, T4 und/oder T5 vollständig oder nahezu identisch. Aufgrund dieser Redundanz kann auf die Übertragung von identischen Teilbandsignalen gleicher Frequenzlage in dem Zusatzsignalpaket T3/T4/T5 verzichtet werden. Decoderseitig werden die punktierten Teilbandsignale des Zusatzsignal-Paketes T3/T4/T5 bei der Rekombination nur in ihrem ursprünglichen Tonkanal substituiert. Fehlt beispielsweise ein Teilbandsignal des Center-Kanals S3 im Zusatzsignal-Paket T3/T4/T5, so muß decoderseitig dafür gesorgt werden, daß aus dem linken oder rechten Signal des Stereo-Basissignal-Paketes Lo/Ro nur in dem Center-Wiedergabekanal S3' substituiert wird, jedoch nicht in den beiden anderen Wiedergabekanälen S4' und S5' (Fig. 4). Darüberhinaus müssen die Center-Anteile in den Stereo-Basissignalen Lo' und Ro' vom Dekombinierer 41 unterdrückt werden, da aufgrund des — fehlenden Center-Kanals die Dematrizierung gemäß Gleichungen (6) bis (8) nicht mehr möglich ist.

— Alias-Verzerrungen

Infolge der durch die Punktierung bedingten, nicht-vollständigen Rekonstruktion des Tonsignals im Decoder findet die Auslöschung der Alias-Verzerrungen bei der Rückfilterung in den Synthesefilterbänken 40 (Fig. 4 und 6) nicht vollständig statt. Diese restlichen Alias-Verzerrungen lassen sich im Encoder vorherbestimmen, so daß ihre Maskierung berechnet werden kann. Bei fehlender Maskierung darf keine Punktierung erfolgen.

— Negative Korrelation

Tritt eine negative Korrelation, d. h., Gegenphasigkeit, zwischen Teilbandsignalen gleicher Frequenzlage auf, so hat dies bei der Matrizierung gemäß Gleichungen (1) und (2) zur Folge, daß in dem Links- oder Rechts-Signal des Stereo-Basissignal-Paketes Lo/Ro Auslöschungen auftreten. Diese Auslöschungen werden bei der Dematrizierung gemäß Gleichungen (6) und (7) rückgängig gemacht. Diese Rückgängigmachung unterbleibt jedoch, wenn infolge einer Punktierung eines der Signale T3, T4 oder T5 nicht übertragen wird. Aus diesem Grunde darf bei der Feststellung einer negativen Korrelation keine Punktierung erfolgen.

2. Gemeinsame Mithörschwelle

Die Bitzuweisung der Teilbandsignale in den einzelnen Tonkanälen wird auf der Basis des Signal/Maskierungsabstandes bei einer Berücksichtigung aller Tonkanäle und aller Teilbänder berechnet. Unter Signal/Maskierungsabstand wird das Verhältnis von maximalem Teilbandsignalpegel zu minimaler Mithörschwelle in einem Teilband verstanden.

Daher ist es notwendig, für jeden der Tonkanäle den maximalen Signalpegel und die minimale Mithörschwelle pro Teilband und pro zeitlichem Block zu bestimmen. Die minimale Mithörschwelle wird in der Analysestufe 35 (Fig. 3 und 5) aus einer der Teilbandfilterung parallelgeschalteten Fourieranalyse, der ein psychoakustisches Modell folgt, berechnet. Diese Parallelkonzeption hat den Vorteil, daß damit eine unzureichende Frequenzzerlegung der Filterbank 30 kompensiert werden kann, da einerseits eine gute Spektralauflösung — gegeben durch die Fourieranalyse — für die Bestimmung der spektralen Mithörschwellen notwendig ist und andererseits eine gute Zeitauflösung — gegeben durch die Filterbank 30 — der einzelnen Tonkanäle gegeben ist.

Durch diese Parallelkonzeption können auch Frequenzen und Pegel von Alias-Komponenten ermittelt werden, was besonders wichtig ist hinsichtlich der Operationen, die durch den Kombiniierer 31 (Fig. 3 und 5) durchgeführt werden. Die Entscheidungen dieser speziellen Signalanalyse in 35 werden in das Steuersignal COS, das den Kombiniierer 31 steuert, miteingefügt.

Zum Beispiel haben Experimente gezeigt, daß für den zeitlichen Block zur Berechnung der Mithörschwelle etwa 24 ms ausreichen, was bei einer Abtastfrequenz der Tonkanäle von 48 kHz 1152 Eingangs-PCM Werten entspricht und einen guten Kompromiß zwischen zeitlicher Folge des Tonsignals und/oder baulicher Komplexität der Analysestufe 35 darstellt.

Die Berechnung der genaueren Frequenzauflösung ist jedoch nur encoderseitig notwendig, nicht jedoch im Decoder, was für eine Massenproduktion von Decodern wichtig ist.

Die Berechnung der gemeinsamen Mithörschwelle und des davon abgeleiteten Signal/Maskierungsabstandes beruht auf folgenden Schritten:

1. Schritt
Berechnung der Fouriertransformation zur Zeit/Frequenztransformation der einzelnen Tonkanalsignale.
2. Schritt
Bestimmung des Signalpegels in jedem Teilband für jedes Tonkanalsignal.
3. Schritt
Bestimmung der Ruheshörschwelle, wobei je nach Qualitätsanforderung ein zusätzlicher Sicherheitspielraum zur Erhöhung der Übertragungsdynamik eingesetzt wird.
4. Schritt
Bestimmung der tonalen, d. h. durch einen reinen Ton repräsentierten Komponenten und der mehr geräuschhaften Anteile eines Tonkanals, um eine Unterscheidung in tonale Maskierer und Rauschmaskierer verfügbar zu haben.
5. Schritt
Reduzierung der tonalen und geräuschhaften Komponenten auf gehörphysiologisch relevante Anteile durch Berücksichtigung der gegenseitigen Verdeckung dieser Komponenten und der Verdeckung

durch die Ruheshörschwelle.

6. Schritt
Berechnung der individuellen, d. h. einzelnen Mithörschwellen der relevanten Anteile der tonalen und geräuschhaften Komponenten.

7. Schritt
Bestimmung der globalen Mithörschwelle pro Tonkanal.

8. Schritt
Bestimmung der minimalen Mithörschwelle in jedem Teilband pro Tonkanal.

9. Schritt
Bestimmung des maximalen Wertes aller minimalen Mithörschwellen in Teilbändern gleicher Frequenzlage der unmatrizierten Tonkanalsignale L, R, C, LS, RS.

10. Schritt
Berücksichtigung des Effektes einer reduzierten Verdeckung bei räumlich getrennten Maskierern ("Masking Level Difference" = MLD) bei der Bestimmung des maximalen Wertes gemäß Schritt 9 im Sinne einer Verringerung des ermittelten maximalen Wertes.

11. Schritt
Berücksichtigung einer Ausdehnung der Abhörzone bei der Bestimmung des maximalen Wertes gemäß Schritt 9 und 10. Da die Berechnung der MLD grundsätzlich nur für einen einzigen Abhörort Gültigkeit hat, wird der nach Schritt 9 und 10 ermittelte Maximalwert der gemeinsamen Mithörschwelle zusätzlich reduziert.

System-Konfiguration B

Diese System-Konfiguration wird repräsentiert zum einen durch eine Datenreduktion bekannter Art für die zweikanaligen stereofonen Signale L und R und zum anderen durch eine davon völlig unabhängige Datenreduktion für die Kanäle S3, S4 und S5. Die Kanäle S3, S4 und S5 können dabei voneinander unabhängige Programminhalte aufweisen, beispielsweise als reine Kommentar-Kanäle genutzt werden; sie können aber auch ein weiteres stereofones Tonsignalspaar und einen unbenutzten oder einen beliebigen anderen Tonkanal umfassen. In dem Kombiniierer 31 erfolgt bei der System-Konfiguration B keine Matrizierung der Tonkanäle, d. h., der Kombiniierer 31 ist funktionslos. In der Analysestufe 35 erfolgt nicht mehr die Berechnung einer gemeinsamen Mithörschwelle für die Kanäle S3 bis S5; vielmehr wird für jeden einzelnen Kanal S3 bis S5 die individuelle Mithörschwelle berechnet, woraus entsprechend unterschiedliche Bit-Zuweisungsinformationen BAL für die Quantisierungsstufen 33 generiert werden. Desweiteren unterbleibt in der Analysestufe 35 eine Datenreduktion nach den Regeln des dynamischen Übersprechens (Punktierung). Besonders vorteilhaft für die Konfiguration B ist der Encoder gemäß Fig. 5, bei welchem der Datenstrom MUX in den Datenstrom MUXo eingemischt wird. Sämtliche fünf Tonkanalsignale, deren Bitfluß jeweils zeitlich variiert, werden in einen gemeinsamen, festen Datenrahmen eingefügt. Da sich im statistischen Mittel die Bitflußschwankungen der fünf Tonkanäle näherungsweise kompensieren, ergibt sich eine verbesserte Ausnutzung der Übertragungs- oder Speicherkapazität des kombinierten Datenstromes MUXo*.

Auf der Decoderseite kann der Dekombiniierer 41 bei der System-Konfiguration B als Kanalselektor und als Kanalmischer eingesetzt werden, je nach Steuerung der

Kontroll-Einheit 45 durch die Wiedergabe-Konfigurationseingabe 46.

Desweiteren erzeugt die Analysestufe 35 des Encoders bei der System-Konfiguration B keine Differenz-Skalenfaktorwerte Δ SCF; anstelle dessen werden die Skalenfaktorwerte SCF der Kanäle S3 bis S5 an den Multiplexer 34 gegebenenfalls in codierter Form weitergeleitet. Encoderseitig werden von der Kontroll-Einheit 45 die Skalenfaktoren SCFo ignoriert.

Patentansprüche

1. Verfahren zum Übertragen oder Speichern digitalisierter, mehrkanaliger Tonsignale, bei dem encoderseitig

- a) jedes Tonsignal durch eine Vielzahl spektraler Teilbandsignale digital dargestellt wird, wobei für jedes Teilbandsignal zeitdiskrete, quantisierte Abtastwerte vorliegen;
- b) die Quantisierung der Abtastwerte in den einzelnen Teilbändern nach Maßgabe der kanalspezifischen Intra-Kanal-Mithörschwelle des menschlichen Gehörs im Sinne einer Irrelevanzreduktion geändert (codiert) wird;
- c) für jedes Teilbandsignal ein Skalenfaktor (SCF) bestimmt wird, welcher den Spitzenwert des Teilbandsignals innerhalb eines bestimmten Zeitintervalls (Blocks) klassiert und welcher zur Normierung des Teilbandsignals vor der Quantisierung benutzt wird;
- d) für die Steuerung der Quantisierung der Abtastwerte jedes Teilbandsignals aus einer Tonsignalanalyse eine Bit-Zuweisungsinformation (BAL) gewonnen wird, welche zusammen mit den ermittelten Skalenfaktoren (SCF) und mit den codierten Teilbandsignalen gespeichert bzw. übertragen werden, und bei dem decoderseitig
- e) die codierten Teilbandsignale nach Maßgabe der Bit-Zuweisungsinformation (BAL) decodiert und die Teilbandsignalpegel nach Maßgabe der Skalenfaktoren (SCF) entnormiert werden;
- f) die decodierten und entnormierten Teilbandsignale jedes Tonkanals zu einem breitbandigen, digitalen Tonsignal zusammengesetzt werden, dadurch gekennzeichnet, daß
- g) encoderseitig die Teilbandsignale verschiedener Tonkanäle, jedoch gleicher Frequenzlage interkanalweise nach Maßgabe eines dynamischen Steuersignals (COS) kombiniert werden, welches durch eine an einem binauralen, psychoakustischen Modell orientierten Tonsignalanalyse von mehreren Tonkanälen gewonnen wird, und
- h) decoderseitig die Teilbandsignale verschiedener Tonkanäle, jedoch gleicher Frequenzlage interkanalweise in Abhängigkeit von einer Steuergröße (COM) dekombiniert werden, welche von dem dynamischen Steuersignal (COS) abgeleitet und mitübertragen bzw. mitgespeichert wird.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß encoderseitig mit Hilfe des Steuersignals (COS) und/oder der Bit-Zuweisungsinformation (BAL) eine Punktierung von Teilbandsignalen erfolgt, bei der die Abtastwerte des betreffenden Teilbandsignals nicht übertragen bzw. gespeichert

werden, und daß decoderseitig die nichtübertragenen bzw. nicht-gespeicherten Abtastwerte in den Teilbändern durch übertragene bzw. gespeicherte Abtastwerte in den Teilbändern anderer Tonkanäle, jedoch gleicher Frequenzlage nach Maßgabe der Steuergröße (COM) substituiert werden, wobei die substituierten Abtastwerte hinsichtlich ihrer Pegel an die ursprünglichen Pegelverhältnisse angepaßt werden.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß für die Pegelanpassung aus den Skalenfaktoren (SCF) Gewichtungsfaktoren (GEW) errechnet werden, mit welchen die substituierten Abtastwerte gewichtet werden.

4. Verfahren nach Anspruch 2 oder 3, dadurch gekennzeichnet, daß die Punktierung in Abhängigkeit von der Erfassung einer Aufeinanderfolge von einem transienten und einem stationären Signalzustand desselben Tonkanals oder benachbarten Tonkanäle erfolgt.

5. Verfahren nach einem der Ansprüche 2 bis 4, dadurch gekennzeichnet, daß die Punktierung in den Teilbändern höherer Frequenzlage häufiger erfolgt als in den Teilbändern tieferer Frequenzlage.

6. Verfahren nach einem der Ansprüche 2 bis 5, dadurch gekennzeichnet, daß die Punktierung in denjenigen Teilbändern häufiger erfolgt, in denen die Signalenergie bezogen auf die Signalenergie des gesamten Tonsignals relativ gering ist.

7. Verfahren nach einem der Ansprüche 2 bis 6, dadurch gekennzeichnet, daß die Punktierung unter Ausnutzung des Trägheitseffektes des menschlichen Gehörs erfolgt.

8. Verfahren nach einem der Ansprüche 2 bis 7, dadurch gekennzeichnet, daß die Punktierung unter Ausnutzung der psychoakustischen Verdeckung von benachbarten Tonkanälen erfolgt.

9. Verfahren nach einem der Ansprüche 2 bis 7, dadurch gekennzeichnet, daß die Punktierung unter Ausnutzung von Redundanz in Teilbändern gleicher Frequenzlage von benachbarten Tonkanälen erfolgt, wobei decoderseitig die punktierten Teilbandsignale bei der Rekombination nur in ihrem ursprünglichen Tonkanal substituiert werden.

10. Verfahren nach einem der Ansprüche 2 bis 9, dadurch gekennzeichnet, daß die Punktierung nur dann erfolgt, wenn encoderseitig festgestellte Alias-Verzerrungen bezüglich der menschlichen Höreigenschaften irrelevant sind.

11. Verfahren nach einem der Ansprüche 2 bis 10, dadurch gekennzeichnet, daß die Punktierung unterbleibt, wenn zwischen Teilbandsignalen gleicher Frequenzlage eine negative Korrelation vorliegt.

12. Verfahren zum Übertragen oder Speichern digitalisierter, mehrkanaliger Tonsignale, bei dem encoderseitig

- a) das Tonsignal jedes Tonsignals durch eine Vielzahl spektraler Teilbandsignale digital dargestellt wird, wobei für jedes Teilbandsignal zeitdiskrete, quantisierte Abtastwerte vorliegen;
- b) die Quantisierung der Abtastwerte in den einzelnen Teilbändern nach Maßgabe der kanalspezifischen Intra-Kanal-Mithörschwelle des menschlichen Gehörs im Sinne einer Irrelevanzreduktion geändert (codiert) wird;
- c) für jedes Teilbandsignal ein Skalenfaktor

(SCF) bestimmt wird, welcher den Spitzenwert des Teilbandsignals innerhalb eines bestimmten Zeitintervalls (Blocks) klassiert und welcher zur Normierung des Teilbandsignals vor der Quantisierung benutzt wird;

d) für die Steuerung der Quantisierung der Abtastwerte jedes Teilbandsignals aus einer Tonsignalanalyse eine Bit-Zuweisungsinformation (BAL) gewonnen wird, welche zusammen mit den ermittelten Skalenfaktoren (SCF) und mit den codierten Teilbandsignalen gespeichert bzw. übertragen werden, und bei dem decoderseitig

e) die codierten Teilbandsignale nach Maßgabe der Bit-Zuweisungsinformation (BAL) decodiert und die Teilbandsignalpegel nach Maßgabe der Skalenfaktoren (SCF) entnormiert werden;

f) die decodierten und entnormierten Teilbandsignale jedes Tonkanals zu einem breitbandigen, digitalen Tonsignal zusammengesetzt werden, dadurch gekennzeichnet, daß für den Fall einer stereofonen Interkorrelation von Tonkanälen (Lo, Ro, C, L, S, R, S) encoderseitig die Quantisierung der Teilbandsignale der stereofon interkorrelierten Tonkanäle nach einer von einer gemeinsamen Mithörschwelle abgeleiteten Steuerinformation (BAL) erfolgt, wobei die gemeinsame Mithörschwelle aus den individuellen Mithörschwellen aller betrachteten Tonkanäle gebildet wird, und daß decoderseitig die Dequantisierung der Teilbandsignale der stereofon interkorrelierten Tonkanäle in Abhängigkeit von der Steuerinformation (BAL) erfolgt, welche mitübertragen bzw. mitgespeichert wird.

13. Verfahren nach Anspruch 12, dadurch gekennzeichnet, daß die gemeinsame Mithörschwelle unter Berücksichtigung des Maskierungsunterschiedes bei stereofoner Wiedergabe berechnet wird.

14. Verfahren zum Übertragen oder Speichern digitalisierter, mehrkanaliger Tonsignale, bei dem encoderseitig

a) das Tonsignal jedes Tonsignals durch eine Vielzahl spektraler Teilbandsignale digital dargestellt wird, wobei für jedes Teilbandsignal zeitdiskrete, quantisierte Abtastwerte vorliegen;

b) die Quantisierung der Abtastwerte in den einzelnen Teilbändern nach Maßgabe der kanalspezifischen Intra-Kanal-Mithörschwelle des menschlichen Gehörs im Sinne einer Irrelevanzreduktion geändert (codiert) wird;

c) für jedes Teilbandsignal ein Skalenfaktor (SCF) bestimmt wird, welcher den Spitzenwert des Teilbandsignals innerhalb eines bestimmten Zeitintervalls (Blocks) klassiert und welcher zur Normierung des Teilbandsignals vor der Quantisierung benutzt wird;

d) für die Steuerung der Quantisierung der Abtastwerte jedes Teilbandsignals aus einer Tonsignalanalyse eine Bit-Zuweisungsinformation (BAL) gewonnen wird, welche zusammen mit den ermittelten Skalenfaktoren (SCF) und mit den codierten Teilbandsignalen gespeichert bzw. übertragen werden, und bei dem decoderseitig

e) die codierten Teilbandsignale nach Maßgabe

be der Bit-Zuweisungsinformation (BAL) decodiert und die Teilbandsignalpegel nach Maßgabe der Skalenfaktoren (SCF) entnormiert werden;

f) die decodierten und entnormierten Teilbandsignale jedes Tonkanals zu einem breitbandigen, digitalen Tonsignal zusammengesetzt werden, dadurch gekennzeichnet, daß encoderseitig die Teilbandsignale verschiedener Tonkanäle, jedoch gleicher Frequenzlage interkanalweise in Abhängigkeit des entsprechenden der Systemkonfiguration gewählten Tondarstellungsformates der Tonkanäle kombiniert werden, und daß decoderseitig die Teilbandsignale verschiedener Tonkanäle, jedoch gleicher Frequenzlage interkanalweise in Abhängigkeit eines Steuersignals (CFG) dekombinert werden, welches von dem gewählten Tondarstellungsformat abgeleitet und mitübertragen bzw. mitgespeichert wird.

Hierzu 7 Seite(n) Zeichnungen

- Leerseite -

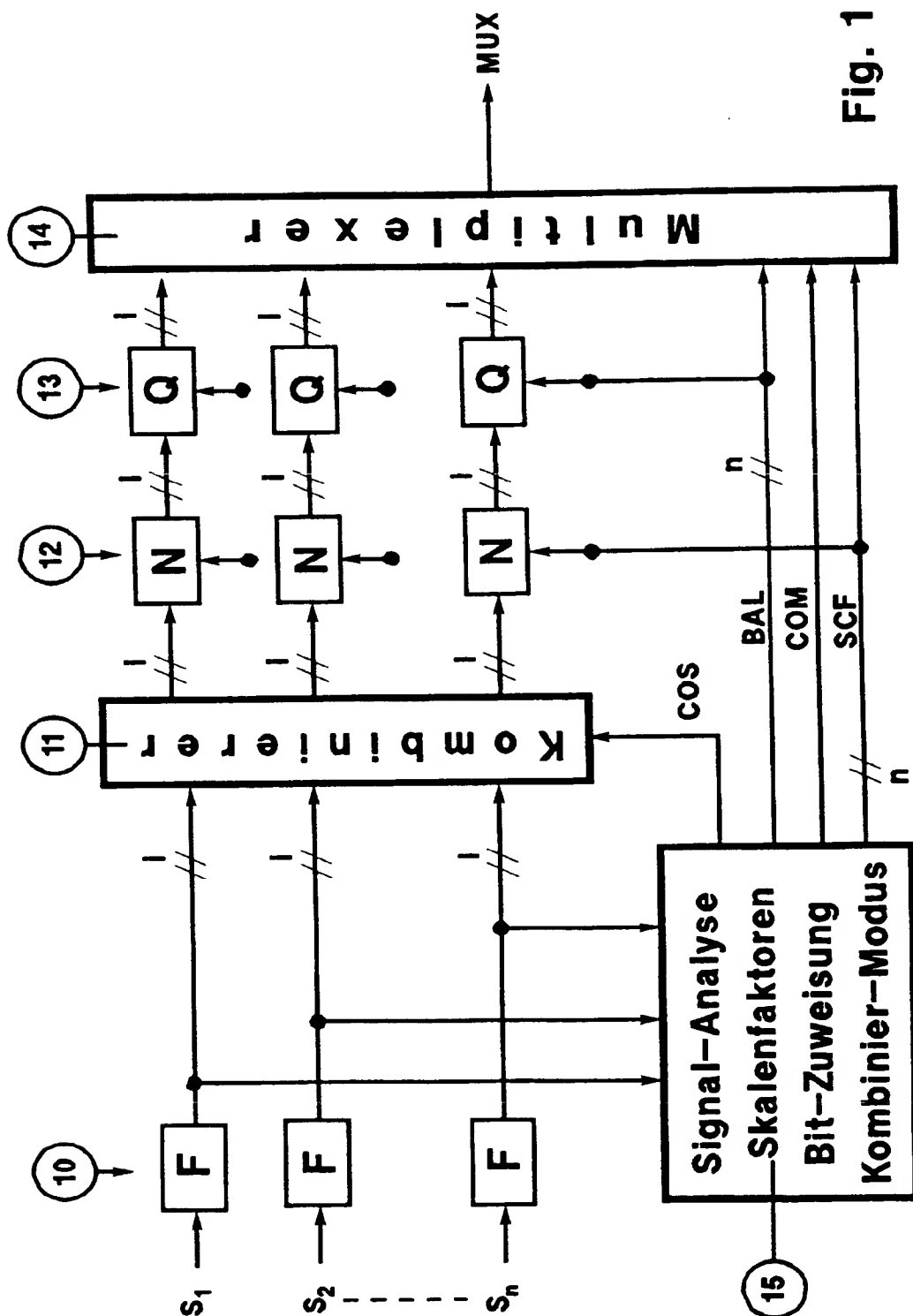
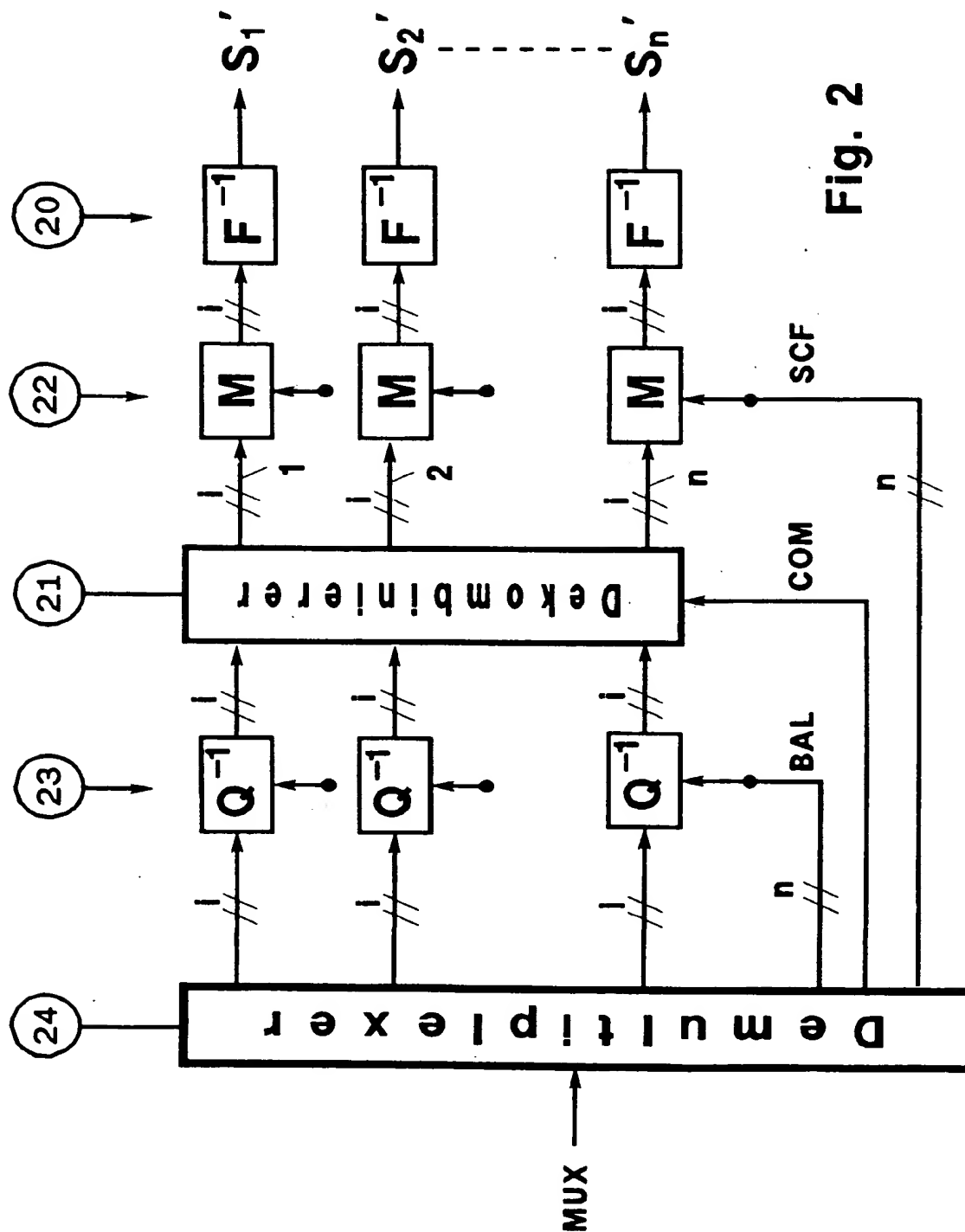


Fig. 1



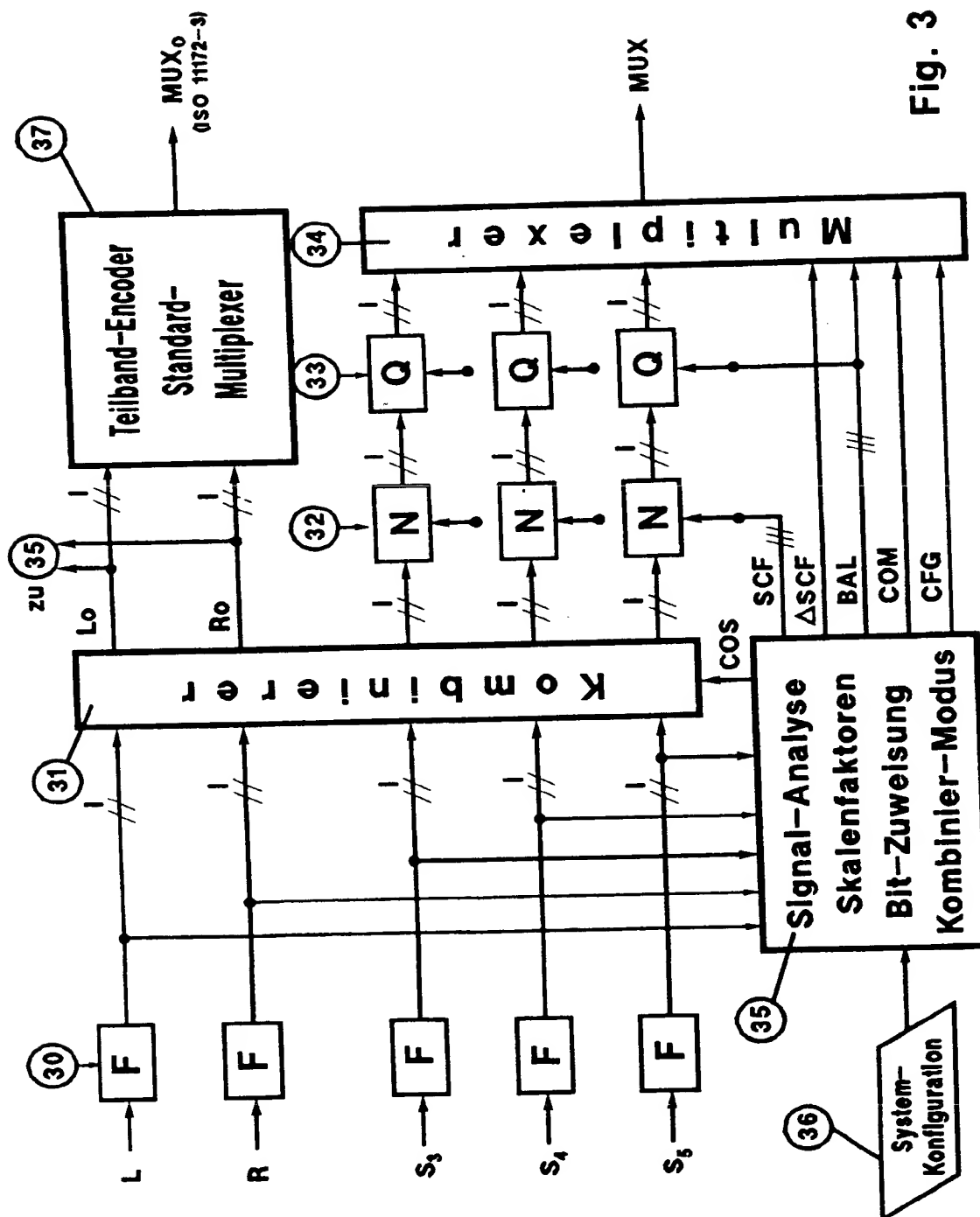
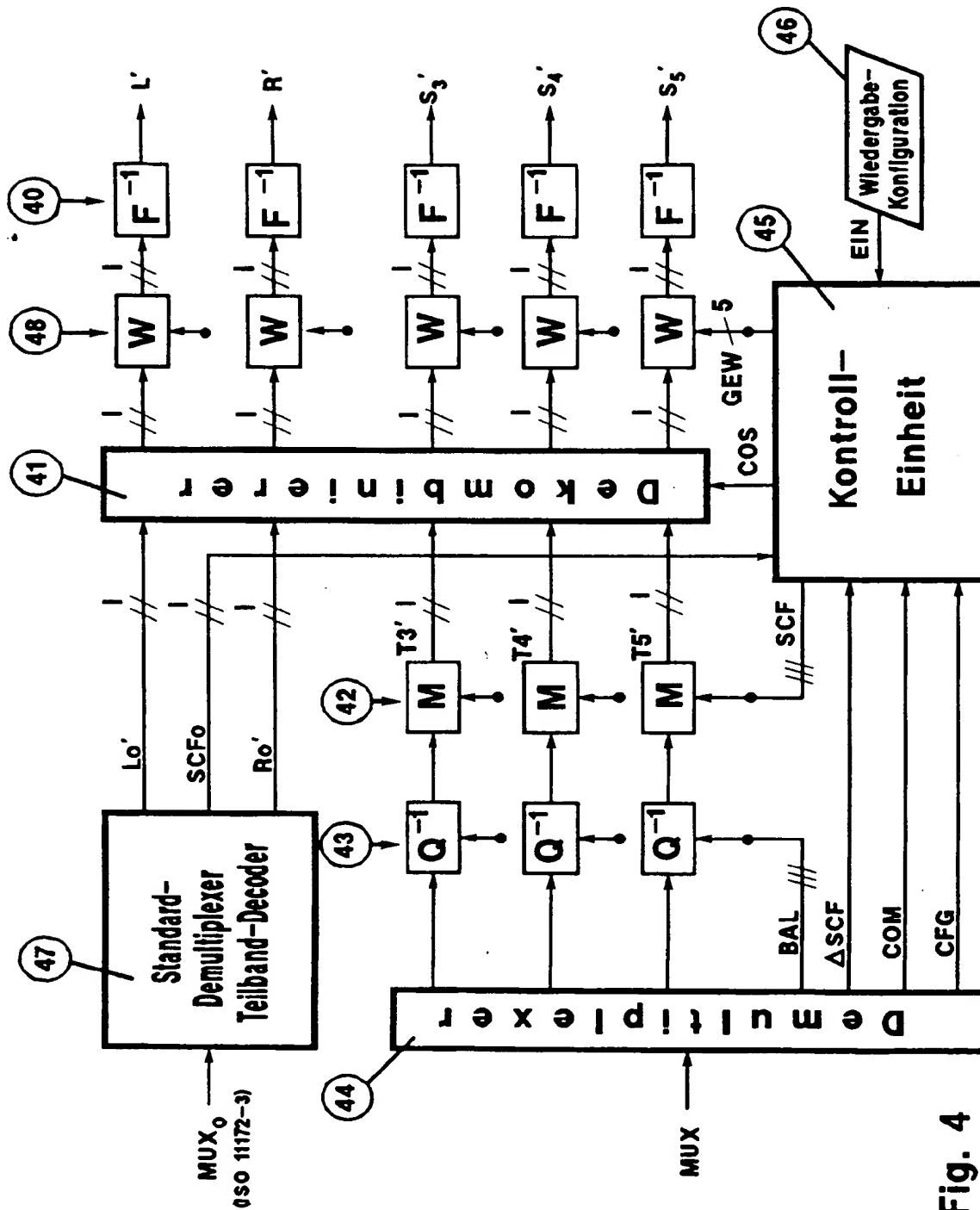


Fig. 3



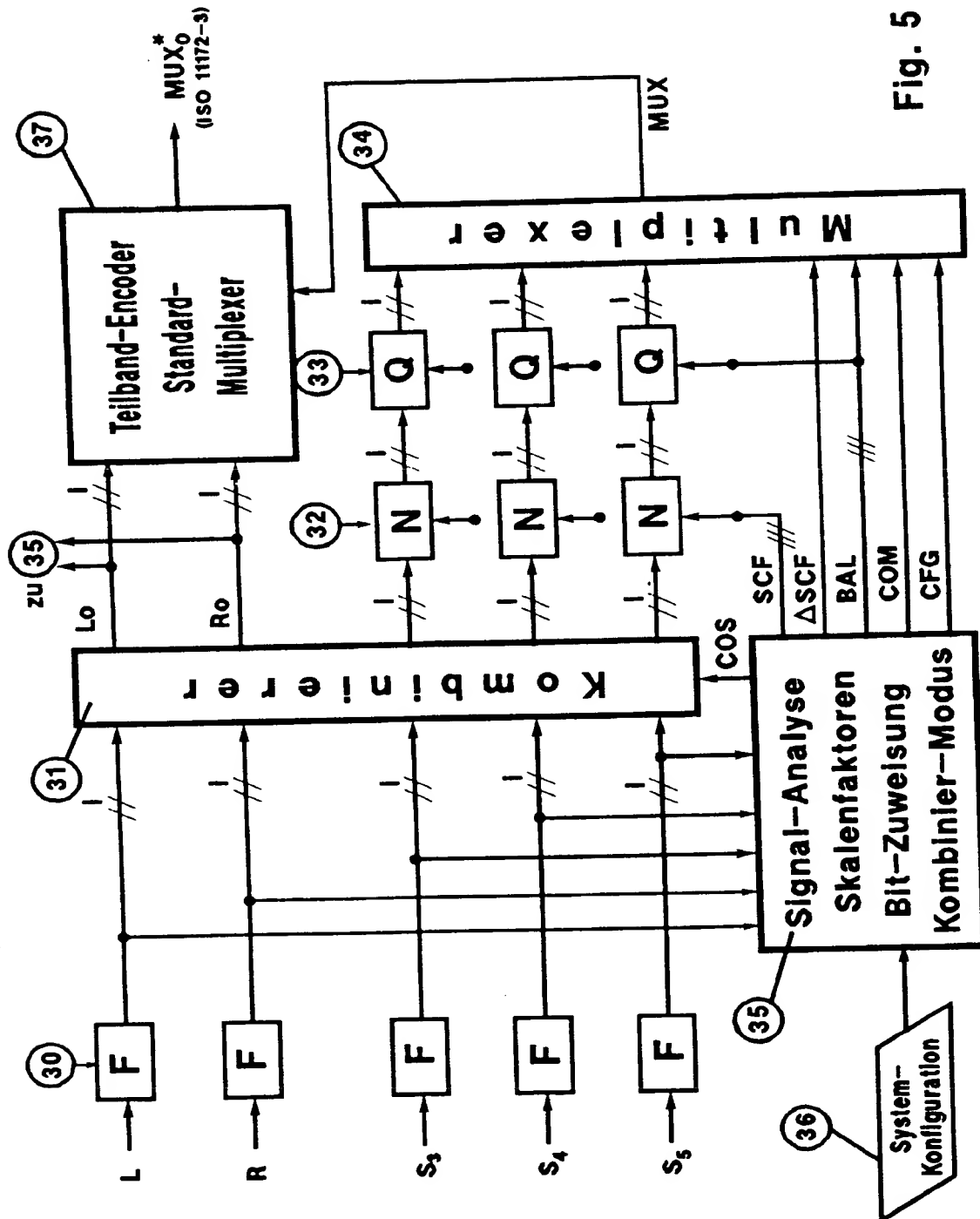


Fig. 5

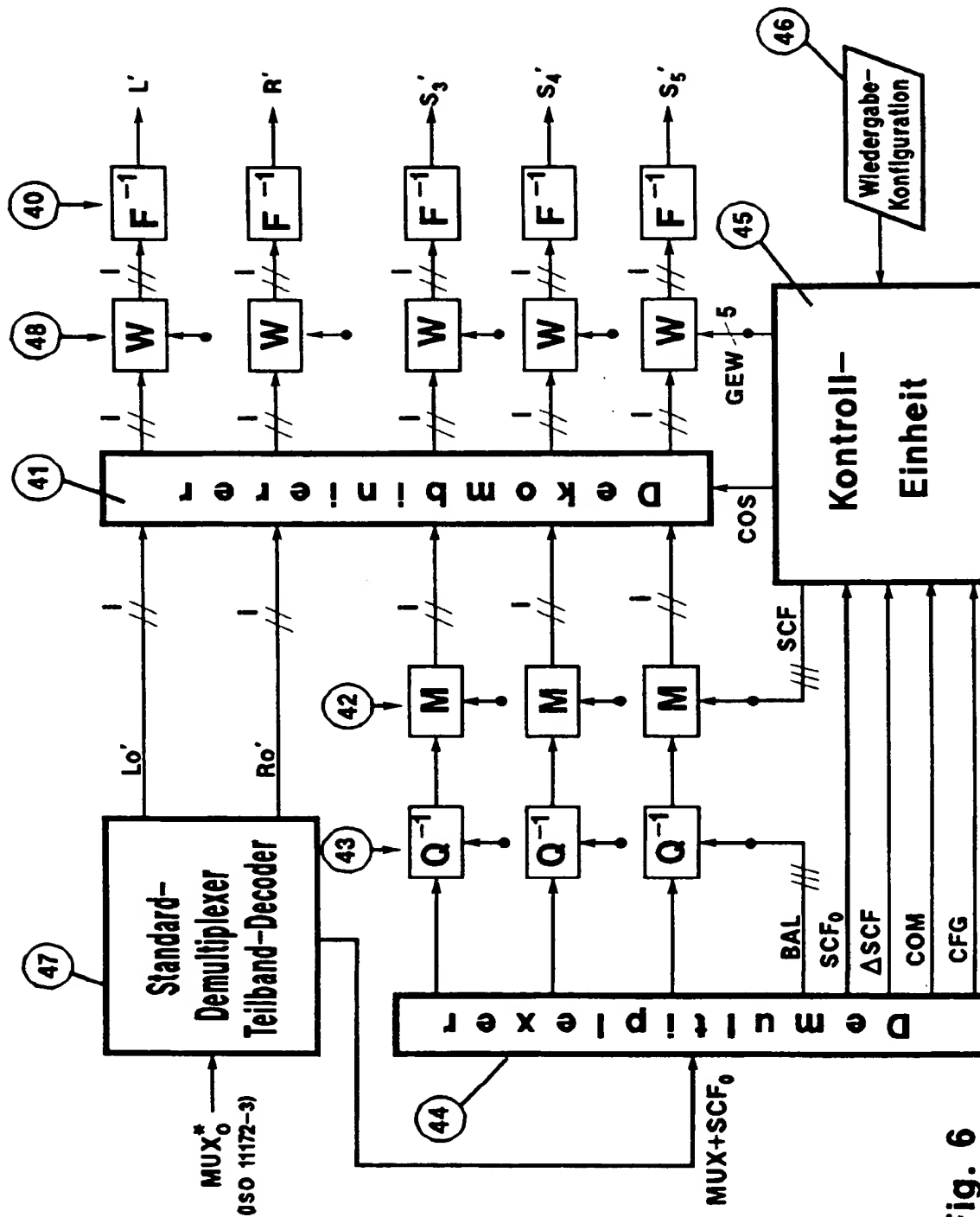


Fig. 6

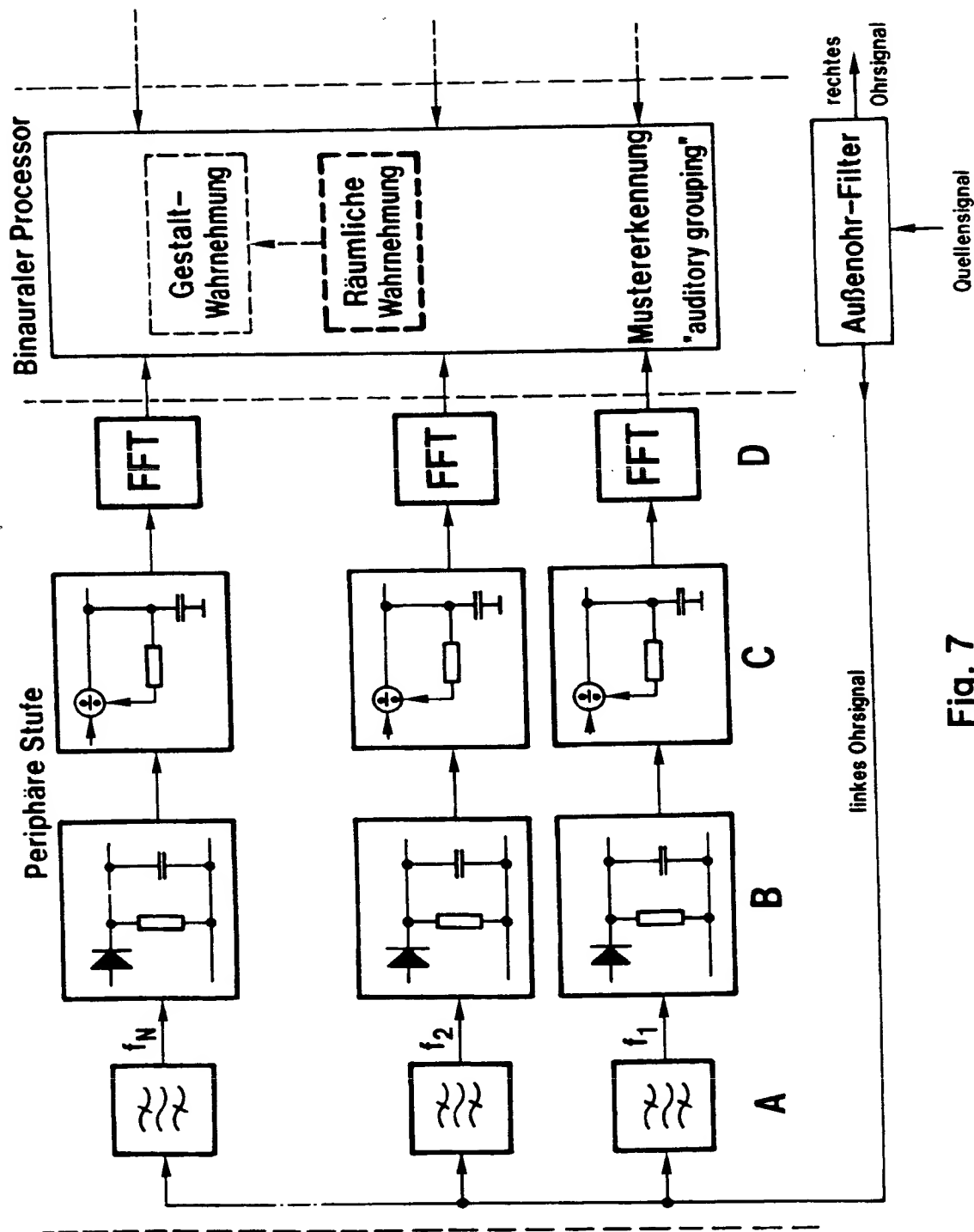


Fig. 7